R E P O R T R E S U M E S

ED 011 734 FL 000 119

EVALUATING THE OBJECTIVES IN FOREIGN-LANGUAGE TEACHING.

BY- VALETTE, REBECCA M.

PUB DATE JUN 66

EDRS PRICE MF-$0.09 HC-$0.52 13P.

DESCRIPTORS- *ACHIEVEMENT TESTS, *MEASUREMENT TECHNIQUES, *MODERN LANGUAGES, *STANDARDIZED TESTS, *TEST CONSTRUCTION, LANGUAGE RESEARCH, AUDIOLINGUAL SKILLS, LANGUAGE TESTS, OBJECTIVE TESTS, LANGUAGE INSTRUCTION KAY SONAGRAPH, MLA COOPERATIVE FOREIGN LANGUAGE TESTS, PIMSLEUR READING COMPREHENSION TESTS

AUDIOLINGUAL TEACHING METHODS HAVE CREATED THE NEED FOR NEW TESTING TECHNIQUES IN THE FOUR LANGUAGE SKILLS OF LISTENING, SPEAKING, AND WRITING. VARIOUS CURRENT METHODS AND SPECIFIC STANDARDIZED TESTS FOR EVALUATING EACH OF THE SKILLS, INCLUDING PURE SKILLS TESTS WHICH UTILIZE ONLY ONE SKILL IN BOTH STIMULUS AND RESPONSE AND HYBRID SKILLS TESTS WHICH EMPLOY TWO OR MORE SKILLS PER ITEM, ARE DESCRIBED. SUGGESTIONS FOR FURTHER RESEARCH AND DEVELOPMENT IN LANGUAGE TESTING INCLUDE THE FOLLOWING--(1) COMPARABLE TESTS IN THE FOUR SKILLS SHOULD BE DEVELOPED, (2) PURE LISTENING COMPREHENSION TESTS SHOULD BE STANDARDIZED, AND THEIR RESULTS NATIONALLY COMPARED TO THOSE OF PURE READING TESTS, (3) THE PRECISE RELATIONSHIP BETWEEN THE COMPREHENSION AND THE PRODUCTION SKILLS SHOULD BE INVESTIGATED, (4) STUDIES FOR EACH FOREIGN LANGUAGE SHOULD DETERMINE THE APPLICABILITY OF "PARTIAL PRODUCTION" TECHNIQUES, AND (5) THE VALIDITY OF THE DICTATION IN MEASURING THE STUDENT'S COMMAND OF THE LANGUAGE SHOULD BE INVESTIGATED. THIS ARTICLE IS PUBLISHED IN THE "INTERNATIONAL REVIEW OF APPLIED LINGUISTICS IN LANGUAGE TEACHING," VOLUME 4, NUMBER 2, 1966. (AUTHOR)

ED011734

1M 21010 F

IRAL

International Review of Applied Linguistics in Language Teaching

Revue Internationale de Linguistique Appliquée: Enseignement des Langues

Internationale Zeitschrift für angewandte Linguistik in der Spracherziehung

---

FL 000 119

IV/2 (1966)

Editors:
A. P. van Teslaar / Paris
H. P. H. Walz / Heidelberg

Editorial Advisory Board:

John B. Carroll, Harvard University · Ahalya Chari, Regional College of Education Mysore · A. Culioli, Université de Paris · Anwar S. Dil, Linguistic Research Group of Pakistan Lahore · Hans Glinz, Universität Zürich und Pädagogische Akademie Kettwig Max Gorosch, University of Stockholm · G. Gougenheim, Université de Paris · Alfred Hayes, Center for Applied Linguistics Washington D. C. · F. Irmen, Universität Heidelberg · W. Lambert, Mc Gill University A. Martinet, Université de Paris · Christine A. E. M. Mohrmann, Secretary of CIPL · P. B. Pandit, University of New Delhi · B. Pattison, University of London · B. Quemada, Université de Besançon · R. Quirk, University of London · P. D. Strevens, University of Leeds

Julius Groos Verlag Heidelberg

Contents – Sommaire – Inhalt

Vol. IV/2 (1966)

RESEARCH PROJECTS AND REPORTS:

M. I. Rocha e Silva/C. B. Ferster: An Experiment in Teaching a Second Language 85

SURVEYS AND GENERAL ARTICLES:

W. v. Raffler Engel: The Natural Basis of Linguistic Levels . . . 115

B. Spolsky: A Psycholinguistic Critique of Programmed Foreign Language Instruction 119

R. M. Valette: Evaluating the Objectives in Foreign-Language Teaching 131

ANNOUNCEMENT: 141

REVIEWS: 143

PUBLICATIONS RECEIVED: 155

Coming issues will include the following articles:

R. Barrutia (Irvine):	Some pedagogical dangers in recent linguistic trends
P. Delattre (Santa Barbara):	A comparison of syllable length conditioning among languages
R. M. Jones (Aberystwyth):	Situational vocabulary
H. Kaluza (Wroclaw):	Proper names and articles in English
H. W. Kirkwood (London):	Translation as a basis for contrastive linguistic analysis
E. A. Levenston (Jerusalem):	A classification of language differences
R. C. Troike (Austin):	Machine translation and second language teaching
A. Valdmann (Bloomington):	Norme pédagogique : Les structures interrogatives du français

© 1963 Julius Groos Verlag

6900 Heidelberg/Germany, Gaisbergstraße 6–8
P. O. Box 629. Cable Address: Groos Heidelberg
Banking Account No. 02/52338 Deutsche Bank Fil. Heidelberg
Printed by Julius Beltz, Weinheim/Germany

Editors: A. P. van Teslaar, SHAPE, Paris
H. P. H. Walz, Heidelberg University, Plöck 50

Published quarterly
Erscheint vierteljährlich

Subscription per Vol.	Europe (excl. France)	Overseas (excl. USA)
Individual	DM 38.– (US$ 9.50)	DM 39.– (US$ 9.75)
Institutional	DM 42.– (US$ 10.50)	DM 43.– (US$ 10.75)
Single Copies	DM 10.50 (US$ 2.65)	DM 11.– (US$ 2.75)

Postage included
France: Price: NF 60 – Sole Agency: Librairie Marcel Didier –
4 et 6, rue de la Sorbonne – Paris V
USA: Price: US$ 10.75 – Sole Agency: Chilton Books, A Division of Chilton Company –
East Washington Square – 525 Locust Street – Philadelphia 6, Penna. (19106)

EVALUATING THE OBJECTIVES IN FOREIGN-LANGUAGE TEACHING *)

Rebecca M. Valette

L'enseignement moderne des langues étrangères met l'accent sur la compréhension et la faculté d'exprimer la langue parlée aussi bien qu'écrite. Cette modification du but à atteindre a entrainé la création de nouvelles formes de test et d'examen destinés à permettre une évaluation rigoureusement objective. Mais il ne suffit pas d'évaluer globalement les connaissances de l'élève en matière de vocabulaire, morphologie et syntaxe; il faut pouvoir évaluer le degré relatif d'habileté de l'élève à comprendre, parler, lire et écrire la langue étrangère. L'auteur examine diverses méthodes d'évaluation, soit pures, soit mixtes, qui seraient spécifiques à chacun de ces domaines et esquisse à la fin quelques directions de recherche et de développement futurs.

Im modernen Fremdsprachenunterricht wird besonderes Gewicht gelegt auf das Verstehen und die Fähigkeit, die gesprochene Sprache so gut zu beherrschen wie die geschriebene. Diese Zielsetzung hatte die Erarbeitung neuer Tests und Prüfungsmethoden zur Folge, die eine möglichst objektive Leistungsbeurteilung ermöglichen sollten. Es ist jedoch nicht ausreichend, das Wissen des Schülers in Wortschatz, Morphologie und Syntax als Ganzes zu beurteilen. Vielmehr muß es möglich sein, beim einzelnen Schüler den relativen Grad dessen herausfinden zu können, wie gut er die Fremdsprache versteht, spricht, liest und schreibt. Vf. untersucht verschiedene Beurteilungsmethoden, reine und gemischte Formen, die für jeden dieser Bereiche speziell geeignet sind und skizziert abschließend einige Anregungen für die Forschung und zukünftige Entwicklung.

Spoken language is a phenomenon exceedingly more complex than its graphic representation, the printed word. However, since the study of literature, which for centuries represented the educational goals of the elite, necessitated the acquisition of a reading knowledge of foreign languages, language and literature were formally equated in the curriculum. In restricting their concept of language, educators disregarded the idea of language as oral communication and chose to ignore the fact that throughout history conquered peoples, immigrants and travellers did adapt to a new linguistic environment. Some people acquired a limited vocabulary, which they modeled according to the grammatical patterns and sound system of their native language. Others achieved proper fluency – in varying degrees, of course. In his new surroundings each individual speaking the language evaluated his own progress in terms of his success in communication. On the other hand, in the schools the scoring of a student's

---

*) Paper read June 28, 1965, in the American Educational Research Association Symposium at the annual meeting of the National Education Association.

ability to listen and to speak was almost totally neglected. Written classroom tests measured the student's grasp of literature.

Now that language teacher have been broadening their aims and developing a curriculum destined not only to teach the core of the language, that is, its words and structures, but also to build up student proficiency in the skills of listening and speaking as well as of reading and writing, new evaluation techniques are being introduced. As in any relatively unexplored field, the initial advances appear spectacular. While in no way detracting from the accomplishments of recent years, I propose to examine the various current methods of measuring the language skills and to suggest areas for further research.

Let us first look briefly at language itself. As you may remember, Alice in Wonderland in her journey through the Looking Glass was intrigued with Jabberwocky:

'Twas brilling, and the slithy toves
Did gyre and gimble in the wabe;
All mimsy were the borogoves,
And the mome raths outgrabe.

The grammatical structure is English: "Slithy" is an adjective modifying "toves" (the plural of "tove"); "gyre" and "gimble" are verbs, as is obvious from their position relative to "did"; "outgrabe" is the past tense of "outgribe". Alice's comprehension problem was uniquely one of vocabulary.

The recently developed modern language curriculum known as the New Key consists of audio-lingual materials which, in Levels I and II, emphasize the acquisition of structure rather than rapid vocabulary development. After all, the structure of the language indicates the relationships among the parts of the sentence. A single unfamiliar word can be looked up in a dictionary, but a new construction poses a dilemma.

Both foreign-language usage and vocabulary adapt themselves well to objective measurement, much as they do in English. The first standardized tests, established in the 1920's, concentrated primarily on these two fundamental aspects of the foreign language; as for English, printed multiple-choice tests proved most reliable and most economical in terms of scoring time.

The new curriculum materials have been devised on the assumption that foreign-language learning is basically a mechanical process of habit formation. Students learn a natural dialogue by ear. The words and structures of that dialogue are presented in a variety of ways, primarily through pattern drills and directed dialogues. Before advancing to the next lesson, the student should be so familiar with the material just covered that he can speak it fluently and correctly. Since New Key proponents define language primarily as "behavior", the material is so structured that the student is induced to speak or behave without making mistakes. The entire lesson, with the exception of perhaps a few minutes

at either the beginning or the end, is conducted in the foreign language. Homework at the elementary level consists in language laboratory sessions or listening at home to tapes or discs recorded by native speakers. Once the student is at ease with the spoken language, he is introduced to the skills of reading and writing, again in such a systematic way as almost to eliminate the possibility of error.

Within the New Key framework, tests conform to rigid qualifications. No wrong forms are employed; consequently, on a multiple-choice test all choices are idiomatic and properly spelled but only one choice constitutes an appropriate answer. Since the foreign language is taught as a system of communication with the least possible reference to English, test items appear entirely in the foreign language. Mixed sentences, partly in English and partly in the target language, are eliminated as inconsistent with the aims. If an English equivalent must be included to clarify the meaning of the item, then two entire sentences should be used. Thus, whereas homework sentences or test items once read:

Je n'aime pas cette robe-ci, je préfère – (that one)

such questions now appear as

I don't like this dress, I prefer that one.

Je n'aime pas cette robe-ci, je préfère –.

In these two sentences the student sees that there is no one-to-one word-to-word correspondence between "*je n'aime pas*" and "I don't like." Thus he has less tendency to look for a French word for "that" and another for "one" but is encouraged to find semantically equivalent structures in the two languages. Finally, New Key examinations include only natural constructions. Contrived sentences containing pitfalls which would stump even the native speaker have been eliminated.

However, the measurement of achievement in language learning must not stop here. Admittedly, the word, the grammatical element and the syntactic pattern are the building blocks of the sentence. Without them one can neither understand nor speak, neither read nor write the language. But since the New Key curriculum emphasizes the acquisition of all four language skills, standardized tests are needed to evaluate the student's relative proficiency in listening, speaking, reading and writing.

A language is not like football, for example. The coach can measure independently the player's knowledge of the vocabulary and the structure of the game, the rules, the specific plays and their code numbers. On the field he can isolate and evaluate proficiency in each skill, in blocking, in passing, in receiving, in punting. But the language instructor cannot teach or test the language skills without employing vocabulary, grammar and syntax. Each language skill, however, has certain unique particularities. Let us look at these distinctive features and see how they can best be evaluated.

The four skills may be measured separately or in combination. In a *pure* test only one skill is utilized. Tests built around two or three skills, for example listening and reading or listening and writing, will be termed *hybrid* tests.

Listening. One problem in aural comprehension is presented by what the linguists call "minimal pairs", words that differ by only one phoneme. An English example may clarify this concept. Many foreigners learning English have trouble distinguishing the vowel sounds in "ham", "hem" and "hymn". A listening test item would read:

Choose the logical rejoinder:
She has got the ham ready.
1. Well, let's sit down to eat.
2. And Mother has offered to sew it for me.
3. So let's sing it through to see how it sounds.

Had the student understood "hem", the completion mentioning sewing would have seemed plausible; had he understood "hymn" he might have imagined a choir rehearsal. For other languages one can prepare similar listening items in which comprehension hinges on the discrimination of a single sound. Both the key sentence and the possible completions or rejoinders would be prerecorded on tape. Another version of the same type of item would have only the key sentence recorded; options 1, 2, and 3 would be pictures showing, respectively, a woman with a ham, a woman holding the hem of a dress, and an organist with a hymnal. It is also possible to record just the key word "ham", and show the student numbered pictures of a ham, a hem, and a hymn. However, a whole sentence furnishes a more natural context.

A second problem unique to the listening skill is the comprehension of rapid conversation. Perhaps you have had the following experience: Lost in Paris, you asked directions of a policeman and found his French response easily comprehensible; but when this same person began talking with another Frenchman, you were totally unable to follow the conversation. Consider for a moment the phrase "jeetjet." It sounds like nonsense syllables. But were a friend to ask "Jeetjet?" you might answer: "No, but I plan to get a sandwich after this symposium." To understand longer sentences, the student must increase his retention span and learn to pick out key words. Recorded listening comprehension tests, presenting a quick dialog followed by clearly enunciated questions and multiple-choice responses, can validly measure the student's ability to understand rapid speech.

Many other types of listening tests exist. Inasmuch as they measure comprehension of a distinctly recorded conversation or passage which is not built around minimal sound distinctions, such tests primarily evaluate student achievement in structure and vocabulary via the listening skill. Even students with poor discrimination and no training in rapid conversation will do well on such examinations if they are familiar with the content of the items.

So far we have discussed pure listening tests, those in which the entire examination is recorded and the student indicates only a letter response on an answer sheet. In the 1930's phonetic accuracy tests were developed in a multiple-

choice format. Since the skills of reading and writing were receiving greater emphasis in the classroom, the student's listening discrimination could be reliably measured in relation to the printed word. Were such a test given in English, the student would read:

1. He's sleeping.
2. He's slipping.
3. He's leaping.

The recorded voice would state once: "He's sleeping". The student is called upon to select the proper phase. In the New Key classroom, where the spoken language is presented before the written language, such items at the elementary level might tend to become spelling tests in that they measure the relationship between the sounds, which are already familiar, and the printed word.

Some hybrid listening tests currently in use are administered with an answer booklet. The student hears a recorded conversation or passage and then answers either spoken or printed questions by indicating his selection among the suggested responses he reads in the booklet. The student who reads with difficulty and is more at ease with the spoken language will be at a disadvantage in this type of "listening" test. The first hybrid listening examinations, employed in the 1920's, avoided this possible danger by presenting the printed section entirely in English.

Speaking. The specific elements characteristic of the speaking skill are pronunciation, intonation and fluency. In 1929, the Modern Foreign Language Study report on achievement tests pessimistically stated: "Standardized group tests for pronunciation and oral composition which could be administered widely seem almost impossibilities"[1]). However, four years later experiments were in progress using phonographic aluminum disks. In recent years the rapid growth of language laboratories has facilitated the administration of identical speaking tests to large groups of students. Trained scorers, often working in groups, have demonstrated the possibility of rating student performance quite objectively. Only specific aspects of each utterance are scored but the student does not know what the examiner will listen for. As early as the 1930's it was noted that reliable scores could be obtained. This was confirmed by an experiment conducted at the University of Colorado in 1960–1962, by the MLA Cooperative Foreign Language Tests, and by the experimental Pimsleur tests.

Unfortunately, most speaking tests are tedious to score because each student tape must be played in its entirety. Since speaking tests are often administered with a spoken cue such as a question or command, the scorers must spend time listening to the same cues on every student tape. Nelson Brooks has

---

[1]) V. A. C. Henmon. *Achievement Tests in the Foreign Languages*, Publications of the American and Canadian Committees on Modern Languages, Vol. 5, New York, 1929, p. 3.

thought of a system whereby the student activates his own tape whenever he speaks and thus records only his own voice [2]).

The ability to read aloud has generally been evaluated on the basis of a relatively long narrative or dramatic selection. George Scherer has found that this type of oral performance can be reliably graded by means of a passage as short as four lines, again reducing the scoring time [3]). Another interesting experiment is underway at the Education Office of the Supreme Headquarters of Allied Powers in Europe (SHAPE) near Paris where pronunciation is being measured with the Kay Sonagraph The student sentence is graphically reproduced so that the relative intensities of frequencies from 0—8000 kilocycles are represented as a function of time. On the sonagram the pronunciation of vowels and consonants can be evaluated visually. A different setting on the machine yields a graphic picture of the intonation. Whereas all voices give a readable intonation curve, it is unfortunately often difficult to read the pronunciation sonagrams of women's voices.

In a pure speaking test the student is asked to talk about a suggested topic, to describe a picture, or to give directions according to a map or diagram. Such tests, scored on fluency and overall quality, are generally administered as the final section of a longer examination.

If students are to be evaluated on their pronunciation of a certain sound or on the intonation pattern they give a particular phrase, then all must utter the same sentence. It appears almost impossible to utilize a pure speaking test to elicit such a response. Consequently other means are employed: the student recites or records a memorized passage or poem; the student repeats a sentence he hears on the tape; the student answers a specific prerecorded question according to a model response; the student reads a printed passage or sentence. In the audio-lingual curriculum, spoken cues or directions seem preferable at the elementary levels. Advanced students more familiar with the printed word may record sentences read from a test booklet. At this point in their training they are less likely to allow the written forms to interfere with their pronunciation. Tests in which students read aloud have the advantage of cutting the scoring time in half since the judges need not listen to spoken cues.

Reading. The reading skill is characterized by speed and the recognition of structure in long or complex sentences. Reading comprehension, in this sense, is similar to reading comprehension in English, an area in which not all American students attain equal proficiency. The speed with which a student reads a foreign language can be measured by timing student performance on a reading comprehension test, or, more objectively, by administering a long test which the

---

[2]) New developments in modern languages testing, *Educational Horizons*, Fall 1964, 21—25.

[3]) George Scherer and Michael Wertheimer. *A psycholinguistic experiment in foreign-language teaching*. New York, 1965, p. 122.

students are unable to finish within a given period of time. Comprehension, in relation to speed, is evaluated by multiple-choice questions based on the text or passage.

When it comes to reading, the advocates of the New Key insist on no more "decoding". If the student is really to *read* the foreign language without mentally translating, new words cannot be introduced too frequently at the early levels.

The widely used standardized tests in foreign languages, the College Boards and the New York Regents, are printed objective tests. Many of the items on these do not evaluate general comprehension, but directly measure grammatical usage and vocabulary. A surprisingly large percentage of items on the MLA Cooperative Tests and the experimental Pimsleur Reading Comprehension tests also hinge directly on vocabulary. This factor might account for some of the coolness which certain teachers have shown the New Key materials. In our test-oriented society, many teachers find themselves judged by their students' performance on such reading comprehension tests. Moreover, many teachers, eager to have their second- and third-year students do well on the College Boards have been supplementing the new audio-lingual materials with outside reading in order to increase students' vocabulary. The teachers also realize that most present standardized listening examinations have written options and that consequently unless the student knows how to read he won't pass a listening comprehension test.

Writing. Unique to the skill of writing are spelling and style. Since free or pure writing tests cannot yet be scored objectively, questions of style must be judged by the individual qualified teacher. Spelling tests, in the form of isolated words or paragraph dictation, depend on a spoken stimulus and are generally limited in use to the classroom.

Forty years ago almost all extramural language tests were written tests. The advent of the standardized language examinations brought with it the introduction of completion items. Soon the recognized ease and economy of mechanical scoring relegated writing tests to the classroom. Currently the MLA Coop tests and the experimental Pimsleur tests are reintroducing writing samples, which generally evaluate the student's active knowledge or recall of foreign language usage and structure. Since objectivity requires that a single specific response be elicited, such tests may assume the form of fill-in-the-blank passages or sentence transformation exercises. The authors of New Key materials, trying to avoid translation by elementary and intermediate students, have popularized two new types of writing tests. The first type is the "dehydrated" sentence. In English, for example, the student would read: "Joan-go-school-yesterday". In the second type of test, the structure retention item, the student is given a model sentence, such as "Joan went to school yesterday" followed by a series of words "girls-come-rehearsal-last-week". The new sentence would read: "The girls came to the rehearsal last week."

Having examined various methods of evaluating the four language skills, let us turn to areas which merit future research.

With the publication of the new MLA Cooperative Foreign Language Tests, standardized examinations are now available in all four language skills. The norms established for these separate tests allow the teacher to appraise each student's standing relative to a nationwide sample. What remains to be done?

1. Comparable tests in the four skills should be developed. Relative student proficiency in the language can only be validly ascertained if equivalent tests in the four skills are created, each test utilizing similar structures and vocabulary. Such tests will permit a more valid evaluation of the relative merits of various teaching techniques.

2. Pure listening comprehension tests should be standardized, and their results nationally compared to those of pure reading tests.

3. Another research project might profitably investigate the precise relationship between the comprehension and the production skills, between recognition and recall. There is a positive correlation between the presently-used listening and speaking tests. Equivalent forms of listening discrimination and pronunciation tests should be devised to establish how high the correlation between hearing and speaking really is under the following conditions:

(a) for students who believe that only the listening test score will determine their grade,
(b) for students who believe that only the speaking test score will determine their grade,
(c) for students who believe that both the speaking and the listening test scores will determine their grade.

Equivalent writing and reading tests could be developed to define with more precision the relationship between recall and recognition in tests of usage and grammatical structure. This battery would similarly be administered to three groups of students who respectively have prepared for a reading test only, a writing test only, and both tests together. A very high correlation in a given category would permit the indirect measurement of student recall through the use of recognition items which permit objective testing, high reliability and economy in scoring.

4. Studies for each foreign language could determine the applicability of what Robert Lado terms the "partial production" technique. In a partial production test, printed items are employed to evaluate the recall skills of speaking and writing. Lado measured pronunciation, stress and intonation in English by having the students indicate similarities between printed words or sentences. For example, which of the following does not clearly rhyme with the others: "food" "good" "stood" "wood"? Such an item is effective in English because of the irregular fit between pronunciation and spelling[4]).

---

4) *Language Testing*. New York, 1964, 95–104.

5. The validity of the dictation in measuring the student's command of the language should be investigated. Various methods of scoring dictations should be compared.

Achievement testing in foreign languages is an open field. Recent advances have been notable, but in the coming years we can look forward to a variety of new developments in language testing.

Rebecca M. Valette
Boston College
Boston, Mass.

"PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY JULIUS GROOS VERLAG HEIDELBERG TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE U.S. OFFICE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER."

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.